Thieme

Classification of Diverticular Disease (CDD) – assessment of the intra- and interobserver agreement in abdominal CT scans

Klassifikation der Divertikelkrankheit (CDD) – Auswertung der Intra- und Interobserver-Übereinstimmung bei abdominalen CT-Scans

Authors
Martina Schmidbauer, Alena Levers, Frank K Wacker, Kristina Imeen Ringe

Affiliations
Institute for Diagnostic and Interventional Radiology, Hannover Medical School, Hannover, Germany

Keywords
diverticulitis, diverticular disease, CDD, agreement, computed tomography

received 10.7.2023
accepted 4.10.2023
published online 8.12.2023

Bibliography
Fortschr Röntgenstr 2024; 196: 591–599
DOI 10.1055/a-2203-3144
ISSN 1438-9029

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Correspondence
Dr. Martina Schmidbauer
Institute for Diagnostic and Interventional Radiology, Hannover Medical School, Carl-Neuberg-Straße 1, 30625 Hannover, Germany
Tel.: +49/5 11/5 32 34 21
schmidbauer.martina@mh-hannover.de

ABSTRACT

Purpose Along with ultrasound, computed tomography (CT) is one of the imaging modalities of choice in patients with suspected diverticular disease (DD). Recently, a newer Classification of Diverticular Disease (CDD) has been proposed. However, its reliability in daily radiological practice has never been proven. Therefore, our aim was to evaluate the intra- and interobserver agreement of the CDD in abdominal CT scans.

Methods In this retrospective study, 481 CT scans of patients with suspected DD were included. Two readers (one board-certified radiologist with 6 years of experience, one 3 rd year radiology resident) individually evaluated all CTs in two reading sessions using the CDD. A composite endpoint of a prior consensus reading, follow-up, and intraoperative findings served as the reference. Intra- and interobserver agreement were calculated using Cohen-k statistic.

Results DD was present in 317 cases (66 %), mostly classified as CDD stage 0, 1b, and 2a (28 %, 30 %, und 14 %). Intraobserver agreement was almost perfect for both readers (kappa 0.93 and 0.88). Interobserver agreement was high and improved from substantial (kappa 0.77) in the first reading session to almost perfect (kappa 0.84) in the second reading session. The interobserver agreement was best for CDD types 0 (diverticulosis) and 2c (free perforated diverticulitis) (mean kappa 0.83 and 0.86) and poorest for CDD types 1a (diverticulitis without phlegmon) and 2b (covered diverticulitis with macroabscess) (mean kappa 0.17 and 0.38). Intra- and interobserver agreement of acute uncomplicated (CDD type 1) and acute complicated diverticulitis (CDD type 2) were substantial to almost perfect (mean kappa 0.63–0.86). Agreement with the reference was almost perfect for both observers (mean kappa 0.86 and 0.82). Administration of rectal contrast did not significantly improve the diagnosis.

Conclusion The CDD is a classification based on relatively clear imaging characteristics, which can be readily applied by radiologists with different expertise. In our study, the CDD had a high intra- and interobserver agreement, enabling a reliable therapy-related categorization of DD.

Key Points

- The Classification of Diverticular Disease (CDD) is an easy-to-use classification for diverticular disease based on relatively clear image features.
- The CDD can be applied equally by radiologists with different levels of experience in the clinical routine.
- The high intra- and interobserver agreement indicates high reliability in the therapy-relevant classification of diverticulitis on CT.

Citation Format

- Schmidbauer M, Levers A, Wacker FK et al. Classification of Diverticular Disease (CDD) – assessment of the intra- and interobserver agreement in abdominal CT scans. Fortschr Röntgenstr 2024; 196: 591–599

ZUSAMMENFASSUNG

Ziel Die Computertomografie (CT) zählt neben dem Ultraschall bei Patienten mit Verdacht auf eine Divertikelkrankheit (DK) zur bildgebenden Modalität der Wahl. Kürzlich wurde

eine neuere Klassifikation der Divertikelkrankheit (CDD) vorgeschlagen, die zunehmend auf bildgebenden Merkmalen basiert. Ziel der vorliegenden Studie war es, die Intra- und Interobserver-Übereinstimmung der CDD bei abdominalen CT-Scans zu bewerten.

Material und Methoden In dieser retrospektiven Studie wurden 481 CT-Scans von Patienten mit Verdacht auf DK eingeschlossen. Zwei Leser (ein Facharzt für Radiologie mit 6 Jahren Erfahrung und ein Assistenzarzt für Radiologie im dritten Weiterbildungsjahr) bewerteten alle CTs einzeln in zwei Lesesitzungen anhand der CDD-Klassifikation. Als Referenzstandard diente ein gemeinsamer Endpunkt aus einer vorausgegangenen, unabhängigen Konsensbewertung, dem klinischen Verlauf und intraoperativen Befunden. Die Intra- und Interobserver-Übereinstimmung wurde anhand der Cohen-k-Statistik berechnet.

Ergebnisse Eine DK lag in 317 Fällen (66 %) vor, darunter am häufigsten die CDD Typen 0, 1b und 2a (28 %, 30 % und 14 %). Die Intraobserver-Übereinstimmung war bei beiden Lesern nahezu perfekt (kappa 0,93 und 0,88). Die Interobserver-Übereinstimmung war hoch und verbesserte sich von beträchtlich (kappa 0,77) in der ersten auf nahezu perfekt (kappa 0,84) in der zweiten Lesesitzung. Die Interobserver-Übereinstimmung war am besten für CDD Typ 0 (Divertikulose) und 2c (freie perforierte Divertikulitis) (mittleres kappa 0,83 und 0,86) und am schlechtesten für CDD Typ 1a (Divertikulitis ohne Phlegmone) und 2b (gedeckt perforierte Divertikulitis mit Makroabszess) (mittleres kappa 0,17 und 0,38). Die Intra- und Interobserver-Übereinstimmung der akuten unkomplizierten (CDD Typ 1) und akuten komplizierten Divertikulitis (CDD Typ 2) war beachtlich bis nahezu perfekt (mittleres kappa 0,63–0,86). Die Übereinstimmung mit der Referenz war für beide Beobachter nahezu perfekt (mittleres kappa 0,86 und 0,82). Die Gabe von rektalem Kontrastmittel führte zu keiner signifikanten Verbesserung der Diagnosik.

Schlussfolgerung Die CDD ist ein Klassifikationssystem, das auf relativ eindeutigen bildgebenden Merkmalen in der CT beruht und von Radiologen mit unterschiedlichen Fachkenntnissen leicht angewendet werden kann. In unserer Studie wies die Klassifikation eine hohe intra- und interobserver-Übereinstimmung auf, die eine zuverlässige therapiebezogene Kategorisierung der DK ermöglicht.

Kernaussagen

- Die Classification of Diverticular Disease (CDD) ist ein leicht anwendbares Klassifikationssystem für die Divertikelerkrankung, das auf relativ eindeutig zu erkennenden Bildmerkmalen beruht.
- Die CDD-Klassifikation kann von Radiologen unterschiedlicher Erfahrung gleichermaßen angewandt werden.
- Die hohe Intra- und Interobserver-Übereinstimmung weist auf eine hohe Verlässlichkeit in der Therapie-relevanten Kategorisierung der Divertikulitis in der CT hin.

Introduction

Diverticular disease (DD) is a common, gradually progressive gastrointestinal disorder with increasing prevalence [1]. DD usually manifests in early adulthood and progresses with advancing age with respect to its anatomical extent and diverticula size. In western countries, up to 30 % of individuals are expected to develop asymptomatic diverticulosis by the age of 50 years and 60–70 % by the age of 80 years [2]. While most people with colonic diverticulosis remain asymptomatic, it is estimated that around 5–20 % will develop symptoms [3, 4]. The clinical spectrum of symptomatic DD ranges from mild abdominal pain up to life-threatening complications including perforation and hemorrhage [5–7].

Diverticulitis is diagnosed based on typical clinical symptoms (e. g., left lower quadrant pain and fever) and elevated blood serum inflammatory parameters. In addition, early radiologic imaging is recommended and used to establish the diagnosis [8]. Besides ultrasound, CT has become a mainstay in patients with suspected diverticulitis due to its excellent sensitivity and specificity [9, 10]. By means of CT, not only confirmation of DD is feasible, but also assessment of disease stage including treatment stratification, and exclusion of important differential diagnoses.

Exact classification of DD is required for stage-related therapy. In the past, various staging systems have been endorsed by different national and international societies [11–16]. Over time, algorithms for DD changed from primarily clinical and surgical to more radiological-based classifications with respect to imaging features, thereby substantiating the increasing impact of imaging. Introduced in 2014 and recently updated in the German S3 guideline for Diverticular Disease and Diverticulitis (2021), the Classification of Diverticular Disease (CDD) is used in German-speaking countries primarily with respect to radiological imaging and diagnosis [8, 17]. The main purpose of this new CDD classification was to establish a more comprehensive and treatment-relevant categorization of separate stages of DD (▶ **Table 1**). Recent studies have shown that the CDD enables reliable staging of disease severity [9, 18].

As with all new staging systems, different aspects of clinical applicability in daily radiological practice have to be examined. Therefore, the aim of our study was to evaluate the intra- and interobserver agreement of the CDD classification in patients undergoing abdominal CT for suspected symptomatic DD.

Materials and Methods

Patients

This retrospective study was approved by the local institutional review board with a waiver of patient consent granted. In our study we used a pre-existing fully characterized patient population in whom imaging was performed for suspected DD from a previous study [9]. A composite endpoint of a consensus reading, intraop-

▶ **Table 1** Classification of Diverticular Disease (CDD).

▶ **Tab. 1** Klassifikation der Divertikelkrankheit (Classification of diverticular disease, CDD).

Term	Synonym	Definition	CDD
Asymptomatic diverticulosis		Identification of diverticula in colon	Type 0
Acute uncomplicated diverticulitis	Diverticulitis without perforation	Diverticulitis without peridiverticulitis	Type 1a
		Diverticulitis with phlegmonous peridiverticulitis	Type 1b
Acute complicated diverticulitis	Diverticulitis with covered perforation	Microabscess (≤ 3 cm), minimal free paracolic air	Type 2a
		Macroabscess (> 3 cm)	Type 2b
	Free perforated diverticulitis	Free air, generalized peritonitis	Type 2c
Chronic diverticular disease	Symptomatic uncomplicated diverticular disease (SUDD)	Typical clinical features	Type 3a
	Relapsing diverticulitis without complications	Recurrent signs of inflammation	Type 3b
	Relapsing diverticulitis with complications	Identification of stenosis, fistulas, conglomerate tumor	Type 3c
Diverticular bleeding		Identification of source of bleeding	Type 4

erative findings, and clinical follow-up from that study served as the reference standard. In total, 481 abdominal CT scans were evaluated. Inclusion criteria were as follows: age ≥ 18 years and surgical or clinical follow-up of at least 4 weeks. Non-diagnostic CT scans were excluded. Abdominal CT scans were acquired on a 64-slice (VCT) or 16-slice (LightSpeed) scanner (both GE HealthCare). Depending on clinical context and contraindications, the CT technique varied, including non-contrast or intravenous contrast-enhanced image acquisition with or without additional oral or rectal contrast application (with rectal contrast (G1): n = 99, without rectal contrast (G2): n = 382).

CT evaluation

Image analysis was performed by two radiologists independently of each other. To address different levels of education, one board-certified radiologist with 6 years of experience (reader A) and one 3rd year radiology resident (reader B) were recruited. Both readers were aware of the clinical indication for imaging and the patients' periods of diverticulitis in order to allow classification of patients into CDD category 3 (chronic DD) but were otherwise blinded to additional clinical data or possible follow-up imaging. In preparation, both radiologists read 20 cases not included in this study together with a senior radiologist in order to get familiar with the classification. Afterwards, both radiologists individually evaluated all CT scans in two reading sessions, separated by a 3-month period in order to minimize recall bias. Image evaluation was performed on a commercially available workstation (Visage 7.1, Pro Medicus Inc) in axial, coronal, and sagittal reformations. If DD was suspected on imaging, findings were classified using the CDD according to the recently updated German S3 guideline for Diverticular Disease and Diverticulitis (2021) [8] (▶ **Table 1**).

Statistical analysis

Statistical analysis was performed using GraphPad Prism 9 (GraphPad Software Inc.). Intraobserver agreement and interobserver agreement between reader A and B in both reading sessions as well as between both readers and the consensus reference standard for CDD stages were calculated using a (weighted) Cohen-k statistic. k-values were interpreted as follows: a value less than 0.20 indicated poor agreement; a value between 0.21 and 0.40 fair agreement; a value between 0.41 and 0.60 moderate agreement; a value between 0.61 and 0.80 substantial agreement; a value between 0.81 and 1.00 almost perfect agreement [19]. To test for potential differences regarding the CDD categories between patients with (G1) and without rectal contrast (G2), the Fisher's exact test was used after exclusion of a Gaussian distribution using Shapiro-Wilk. Being dependent on the (observational) prevalence of the characteristic, comparison of different patient cohorts based on Cohen's kappa has only very limited validity [20]. Therefore, agreement between G1 and G2 was given as a percentage, and evaluation of the interobserver agreement and the agreement with the reference standard was based on the second reading session. For all measurements, $p < 0.05$ indicated a significant difference.

Results

Based on consensus reading and clinical or histological/surgical findings as the reference, DD of the colon was present in 317 (66 %) cases. The frequency of the categories according to the CDD classification is given in ▶ **Table 2**. DD was mostly classified as CDD stage 0 in 28 % of cases (n = 88), stage 1b in 30 % (n = 97), and stage 2a in 14 % (n = 45). Of all 481 CT scans, DD was diagnosed by reader A in 335 cases (70 %) and by reader B in 357 cases (74 %). DD was predominantly diagnosed as stage 0 in 29 % and

▶ **Table 2** Frequency of diverticular disease (DD) categories according to the CDD.

▶ **Tab. 2** Häufigkeit der Kategorien der Divertikelkrankheit (DK) gemäß der CDD-Klassifikation.

CDD classification	Reference (G1/G2)	Reader A	Reader B
DD	317 (71/246)	335	357
0	88 (4/74)	96	88
1*	102	65	104
1a	5 (0/5)	4	3
1b	97 (19/78)	61	101
2*	106	96	79
2a	45 (14/31)	62	38
2b	31 (8/23)	30	25
2c	28 (8/20)	24	26
3a	0	0	0
3b	19 (6/13)	19	19
3c	4 (2/2)	0	0
4	0	0	0

G1: with rectal contrast; G2: without rectal contrast.
* Subgroups of CDD types 1 and 2 were combined to a common category.

25 % of cases, stage 1b in 18 % and 28 % and stage 2a in 19 % and 11 %, respectively. With regard to the CDD categories, there were no statistically significant differences between patients who received rectal contrast (G1) and those who did not (G2). However, there were more severe cases in the subgroup with rectal contrast agent (CDD type 0: 19 % vs. 30 %, CDD type 1: 26 % vs. 34 %, CDD type 2: 44 % vs. 30 %, CDD type 3: 11 % vs. 6 %).

Intraobserver agreement

Intraobserver agreement was almost perfect for both readers (reader A: 88.4 % agreement, weighted kappa 0.93; reader B: 84.0 % agreement, weighted kappa 0.88) (▶ **Table 3**). Disagreement was observed in n = 47 cases for Reader A and in n = 57 cases for reader B, mostly related in CDD type 1b/2a (n = 16 and n = 14) and 1b/2b for reader A (n = 11) and types 2a/2b and 1a/0 for reader B (n = 9 each). Discrepancy in DD severity between the two reading sessions was slightly pronounced in the resident compared to the board-certified radiologist. At subgroup analysis, agreement was substantial to almost perfect for all CDD stages and both readers (mean kappa 0.73–1.00) except for CDD type 1a (mean kappa 0.49, indicating moderate agreement). Of note, for CDD stage 1a there was a significant difference between the intraobserver agreement of the two readers (substantial versus fair agreement). Inconsistency between CDD category 1 (acute uncomplicated diverticulosis) and CDD category 2 (acute complicated diverticulosis) was similar among the two observers (n = 27 for reader A and n = 20 for reader B). When combining subgroups CDD types 1 and 2 in a single category, intraobserver agreement was substantial for category 1 (mean kappa 0.77) and almost perfect for category 2 (mean kappa 0.83) for both readers.

The mean intraobserver agreement between the subgroups G1 and G2 was in total at a comparably very high level (86 % and 88 % agreement). Except for CDD type 1a, which seems to be negligible given only a small number of cases (G1 mean n = 2, G2 mean n = 5), agreement was moderately better only for CDD type 1b in patients without rectal contrast (77 % vs. 64 %).

Interobserver agreement

Interobserver agreement was substantial in reading session 1 (68.8 % agreement, kappa 0.77) and almost perfect in reading session 2 (75.6 % agreement, kappa 0.84) (▶ **Table 4**). Except for CDD type 1a, an improvement from the first to the second reading session could be observed for each subtype: significant improvement for CDD type 1b (moderate vs. substantial agreement, kappa 0.58 to 0.70) and type 2b (fair vs. moderate agreement, kappa 0.29 to 0.47).

Classification according to the CDD by the two readers A and B was discrepant in 143 cases (30 %) in the first reading session and in 115 cases (24 %) in the second reading session. In 86 cases, the disagreement was identical in both readings. In n = 41 (first reading session) and n = 34 (second reading session), there was a discrepancy regarding whether diverticulosis was present or not. In general, there was a tendency towards higher CDD categories for reader A compared to reader B, mostly between CDD type 1b/2a (first/second reading n = 28/24) and 2a/2b (n = 28/26). Representative examples of differently classified cases of DD by readers A and B are presented in ▶ **Fig. 1**. In the subgroup analysis, the level of agreement was almost perfect for CDD types 0, 2c, and 3b (mean kappa: 0.83, 0.86, and 0.99, respectively) and substantial for type 1b (mean kappa: 0.64). Less consensus was observed for CDD types 2a and 2b (moderate and fair agreement, mean kappa: 0.47 and 0.38, respectively). The poorest agreement was seen in CDD type 1a (slight agreement, kappa 0.17). When combining subgroups in CDD type 1 (acute uncomplicated diverticulitis) and 2 (acute complicated diverticulitis) in a single category, interobserver agreement was substantial and could be improved from the first to the second reading session in both groups (category 1: mean kappa 0.63, category 2: mean kappa 0.72).

The mean interobserver agreement in the G1 and G2 subgroups was in total equally high (75 % and 74 % agreement). With rectal contrast, agreement between the two observers was higher for complicated diverticulitis, especially for covered perforated diverticulitis (CDD type 2a: 48 % vs. 32 %; CDD type 2b: 62 % vs. 28 %). However, the percentage of discrepant classifications concerning CDD types 1 and 2 was not significantly different with or without rectal contrast (G1: 38 % vs. G2: 42 % of all disagreements).

Acute complicated diverticulitis without free perforation (CDD type 2a/b)

In addition to the German S3 guideline for Diverticular Disease and Diverticulitis recently updated in 2021 [8], we performed a

▶ **Table 3** Intraobserver agreement of diverticular disease (DD) stages according to the CDD.

▶ **Tab. 3** Intraobserver-Übereinstimmung der Stadien der Divertikelkrankheit (DK) gemäß der CDD-Klassifikation.

CDD	Reader A	Reader B	Mean	G1	G2
	Kappa (95 % CI)	Kappa (95 % CI)	Kappa	Mean agreement (%)	Mean agreement (%)
Overall	0.93	0.88	0.91	86	88
0	0.99 (0.97–1.00)	0.85 (0.79–0.91)	0.92	85	87
1*	0.77 (0.69–0.85)	0.77 (0.71–0.84)	0.77	86	77
1a	0.72 (0.42–1.0)	0.26 (0.01–0.52)	0.49	50	23
1b	0.75 (0.66–0.83)	0.80 (0.73–0.86)	0.78	64	77
2*	0.82 (0.76–0.88)	0.84 (0.78–0.90)	0.83	89	85
2a	0.80 (0.72–0.88)	0.71 (0.61–0.82)	0.76	71	72
2b	0.73 (0.61–0.84)	0.72 (0.59–0.85)	0.73	70	67
2c	0.98 (0.94–0.10)	0.94 (0.87–1.00)	0.96	95	92
3b	1.00 (1.00–1.00)	0.97 (0.92–1.00)	0.99	100	100

G1: with rectal contrast; G2: without rectal contrast.
* Subgroups of CDD types 1 and 2 were combined to a common category.

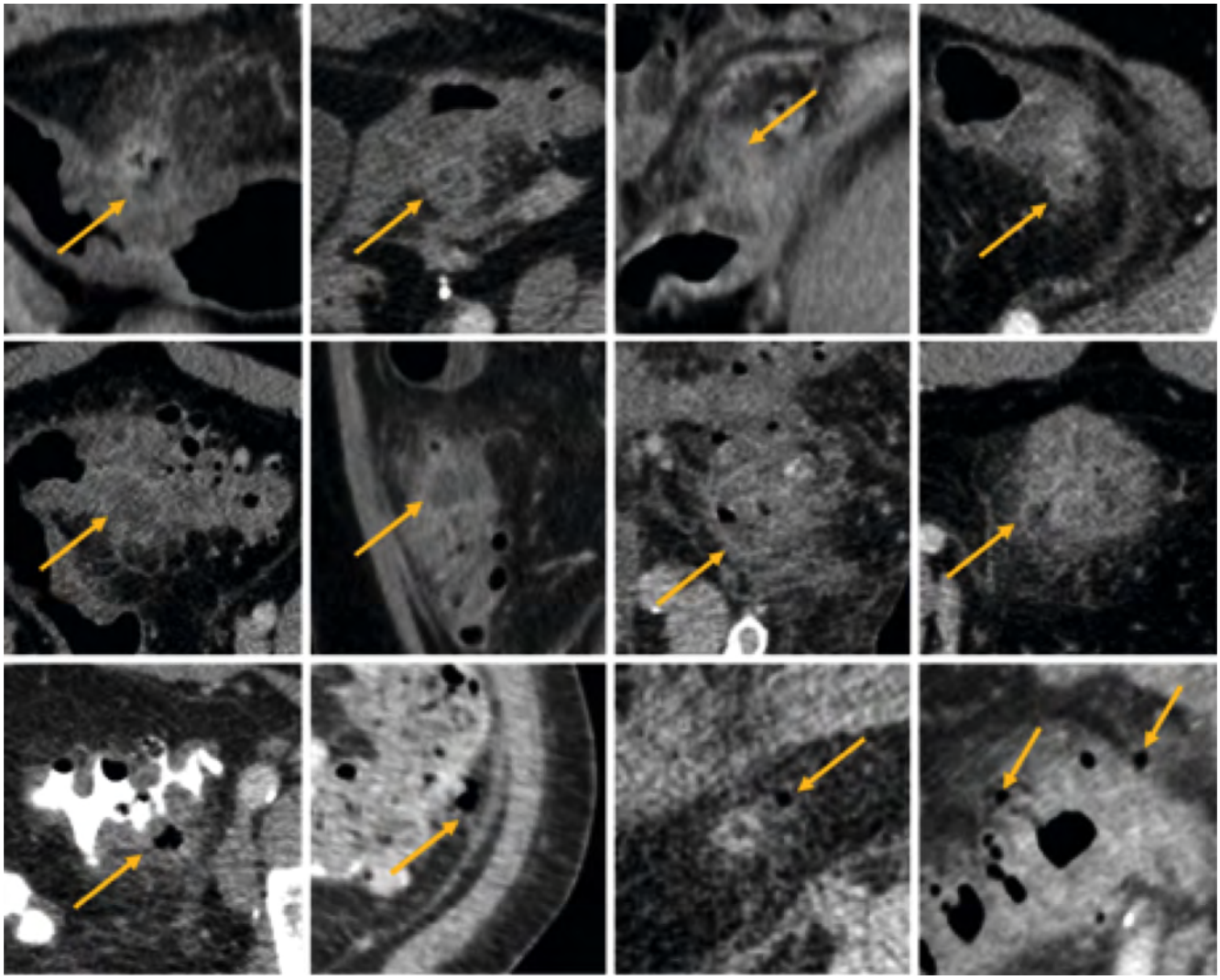
▶ **Table 4** Interobserver agreement of diverticular disease (DD) stages according to the CDD.

▶ **Tab. 4** Interobserver-Übereinstimmung der Stadien der Divertikelkrankheit (DK) gemäß der CDD-Klassifikation.

CDD	Session 1	Session 2		G1	G2
	Kappa (95 % CI)	Kappa (95 % CI)	Mean Kappa	Agreement (%)	Agreement (%)
Overall	0.77	0.84	0.81	75	74
0	0.81 (0.74–0.88)	0.84 (0.78–0.91)	0.83	69	81
1*	0.55 (0.45–0.64)	0.67 (0.58–0.75)	0.63	52	64
1a	0.23 (–0.05–0.52)	0.11 (–0.12–0.33)	0.17	25	0
1b	0.58 (0.48–0.67)	0.70 (0.62–0.79)	0.64	59	66
2*	0.68 (0.60–0.76)	0.75 (0.68–0.83)	0.72	80	69
2a	0.43 (0.31–0.55)	0.50 (0.39–0.62)	0.47	48	32
2b	0.29 (0.13–0.44)	0.47 (0.32–0.63)	0.38	62	28
2c	0.83 (0.71–0.94)	0.88 (0.78–0.97)	0.86	70	84
3b	0.97 (0.92–1.00)	1.00 (1.00–1.00)	0.99	100	100

G1: with rectal contrast; G2: without rectal contrast.
* Subgroups of CDD types 1 and 2 were combined to a common category.

modified analysis based on the initially proposed classification as presented in the prior S2k guideline [17]. CDD types 2a and 2b represent covered perforated stages of diverticulitis and are defined by the abscess size. In the newer S3 guideline, the abscess size threshold has been increased from 1 cm to 3 cm to differentiate between a micro- and a macroabscess. Applying the revised definition with a threshold of 3 cm resulted in downstaging from CDD type 2b to CDD type 2a in 46–60 % of cases (reader A: n = 89 to n = 34, reader B: n = 51 to n = 28). For both sizes, intraobserver agreement was substantial (kappa 0.74 and 0.78, respectively) with improvement for both readers compared to a 1 cm threshold (reader A: kappa 0.80, 95 % CI [0.66–0.94] vs. kappa 0.84, 95 %, CI [0.73–0.96]; reader B: kappa 0.67, 95 % CI [0.49–0.86] vs. kappa 0.72, 95 % CI [0.58–0.90]). The interobserver agreement for CDD



▶ **Fig. 1** CT image examples of interobserver disagreement. Different examples of interobserver disagreement when applying the CDD classification are presented. Interpretation of inflamed diverticulum vs. covered perforation (upper row), highly edematous intestinal wall vs. perforated wall with mural abscess (middle row), and gas-filled diverticulum vs. small neighboring free air bubble (lower row) can be challenging in some cases.

▶ **Abb. 1** Bildbeispiele der Interobserver-Unstimmigkeit in der Kategorisierung der Divertikelerkrankung. Gezeigt sind verschiedene Fallbeispiele der Divertikelkrankheit, die von den Lesern unterschiedlich kategorisiert wurden. Die Abgrenzung eines entzündeten Divertikels gegenüber einer gedeckten Perforation (obere Reihe), einer stark ödematösen Darmwand im Vergleich zu einer perforierten Wand mit intramuraler Abszedierung (mittlere Reihe), oder eines luftgefüllten Divertikels gegenüber einer kleinen benachbarten freien Luftblase (untere Reihe) kann in einigen Fällen schwierig sein.

types 2a and 2b was poor for both the 1 cm and 3 cm threshold (fair agreement, mean kappa: 0.30 and 0.27, respectively).

Agreement between observers and reference standard

In total, consensus in CDD categorization between the reference standard and reader A was observed in n = 247 and 261 (mean agreement 81 %) and reader B in n = 229 and 239 (mean agreement 76 %), respectively. For both readers almost perfect agreement was observed (mean kappa: 0.86 and 0.82, respectively) (▶ **Table 5**). For reader B agreement improved from the first to the second reading session from substantial to almost perfect. Poor consensus was observed for CDD type 1a for both readers and limited agreement was observed for reader A for CDD type 1b (mean 65 %) and for reader B in CDD types 2a and 2b (mean 48 % and 57 %).

Consensus in CDD categorization was high, with or without rectal contrast (G1: mean agreement 75 %, G2: mean agreement 80 %). In the group with rectal contrast, better agreement could be observed for complicated diverticulitis (CDD category 2) compared to patients without rectal contrast (93 % vs. 67 %). The percentage of discrepant classifications concerning CDD types 1 and 2 was not significantly different (G1: 48 % vs. G2: 54 % of all disagreements).

▶ **Table 5** Agreement between observers and the reference standard.

▶ **Tab. 5** Übereinstimmung mit dem Referenzstandard.

		0	1a	1b	2a	2b	2c	3b	3c	Kappa
	Reference	**88**	**5**	**97**	**45**	**31**	**28**	**19**	**4**	
Session 1	Reader A	82 (93 %)	1 (20 %)	61 (63 %)	36 (80 %)	24 (77 %)	24 (86 %)	18 (95 %)	0 (0 %)	0.85 (80 %)
	Reader B	74 (84 %)	0 (0 %)	75 (77 %)	24 (53 %)	16 (52 %)	23 (82 %)	18 (95 %)	0 (0 %)	0.80 (76 %)
	Mean	89 %	10 %	70 %	67 %	65 %	84 %	95 %	0 %	
Session 2	Reader A	83 (94 %)	1 (20 %)	64 (66 %)	41 (91 %)	28 (90 %)	25 (89 %)	18 (95 %)	0 (0 %)	0.87 (84 %)
	Reader B	81 (92 %)	0 (0 %)	76 (78 %)	19 (42 %)	19 (61 %)	25 (89 %)	18 (95 %)	0 (0 %)	0.83 (76 %)
	Mean	93 %	10 %	72 %	67 %	76 %	89 %	95 %	0 %	
G1	Reference	13	0	18	14	8	8	6	2	
	Mean	8 (92 %)		11 (61 %)	11 (78 %)	8 (100 %)	8 (100 %)	6 (100 %)	0 (0 %)	0.79 (75 %)
G2	Reference	74	5	78	31	23	20	13	2	
	Mean	69 (93 %)	1 (20 %)	59 (76 %)	21 (68 %)	11 (47 %)	17 (85 %)	12 (92 %)	0 (0 %)	0.84 (80 %)

Absolute numbers of agreement and percentage of reference for each CDD category are given.G1: with rectal contrast; G2: without rectal contrast.

Discussion

The Classification of Diverticular Disease is based on radiological findings, which are linked to different, more unified treatment options, when compared with previously published systems. Before any classification can be applied in practice, its reproducibility should be put to the test and shown to be as robust as possible. To the best of our knowledge, this is the first study evaluating the intra- and interobserver agreement of the CDD classification in detail since its initial publication in 2014. In the present study, we demonstrated that the CDD may be utilized with high intra- and interobserver agreement, independent of the level of expertise of the radiologist reading the scans.

As mentioned in the introduction, there are different classifications for DD staging. Ünlü et al. analyzed the interobserver agreement of CT stages of diverticulitis according to the modified Hinchey, the Ambrosetti, and the Dharmarajan classification, which are applied especially in the Anglo-American region [11–13]. The authors demonstrated a median overall interobserver agreement with kappa values between 0.72 and 0.83 [21]. This is in line with our results assessing DD using the CDD classification, which show an overall substantial to almost perfect interobserver agreement, even in the case of different levels of experience. Interobserver agreement improved from the first reading session to the second. Hence, we estimate that radiologist training and level of experience may further improve the applicability of the CDD. This might also be an explanation for the slightly lower intraobserver agreement of the less experienced reader B compared to reader A, a board-certified radiologist. Nevertheless and more importantly, even with less experience when first using the CDD classification, substantial agreement with the reference could be achieved, similar to that of the board-certified radiologist.

When analyzing the subtypes of the CDD categories, both intra- and interobserver agreement was highest in CDD types 0 and 2c. Diverticula (type 0) and free abdominal air (type 2c) are radiological features that can be easily detected on CT scans. However, in our study in a moderate number of cases there was an interobserver discrepancy regarding whether or not diverticulosis was present. In almost all of these cases only a marginal number of small diverticula were present. This seems negligible, as asymptomatic diverticulosis does not require treatment. In contrast to these almost clear, objective radiological features, visualization of wall thickening (CDD type 1a) may depend on dilatation of the bowel and is interpreted in a more subjective way, especially in less severe forms. In addition, it can be challenging to determine if small air bubbles next to an air-filled diverticulum are outside or inside the intestinal lumen (CDD type 2a). The interpretation of these findings may be more dependent on the radiologist's experience and could explain the poor interobserver agreement for the classification of CDD subtypes 1a and 2a. However, by combining the subtypes of CDD category 1 and category 2, acute uncomplicated diverticulitis and acute complicated diverticulitis could be easily discriminated from each other with high intra- as well as interobserver agreement. This is important because both categories are related to different therapy strategies: acute uncomplicated diverticulitis (CDD type 1) can be treated primary conservatively and potentially on an outpatient basis, whereas patients with acute complicated diverticulitis (CDD type 2) are usually managed on an inpatient basis, possibly including surgical or interventional treatment [8, 22].

As mentioned above, considering the reduced agreement of reader B (resident) with the reference for CDD type 2, less obvious perforations and abscesses are probably not as easy to diagnose as assumed. It has been postulated that enteric contrast administration helps to distinguish intraluminal from extraluminal air and

fluid collections [23]. This might be underlined by the tendency for better agreement in patients with rectal contrast for perforated diverticulitis. However, the percentage of discrepant classifications between CDD types 1 and 2 was similar and thus the therapy-relevant differentiation between uncomplicated and complicated diverticulitis could not be significantly improved with rectal contrast agent. This could be due to the fact that the absence of extraintestinal findings does not exclude the possibility of a (covered) alimentary tract perforation [24].

In the recently updated S3 guideline on CDD, an enlarged abscess size of 3 cm was determined as a new threshold to discriminate between CDD stage 2a (microabscess) and 2b (macroabscess). Our study showed that intraobserver agreement could be improved using the 3 cm abscess size threshold. However, interobserver agreement was only fair for both thresholds. In contrast to CDD categories 1 and 2, until now there is no sufficient evidence for practical distinction between CDD subtypes 2a and 2b. In the recently published bicentric observation study VADIS (Validation of the German Classification of Diverticular Disease), Lauscher et al. validated the CDD classification and detected a difference between patients with micro- and macroabscess regarding quality of life and the need for surgery [25]. While patients with CDD type 2a could be treated conservatively on a long term basis, all patients with CDD type 2b required surgery within 2 years of follow-up. Although the redefined abscess size of 3 cm does not significantly improve the practical applicability of the CDD classification from the radiological point of view, it reflects the practical approach more accurately, e. g., therapeutic options like percutaneous drainage. Furthermore, when correlating imaging findings in patients with CDD type 2b with intraoperative findings, we found a slight tendency toward overstaging on CT, when the 1 cm threshold was applied [9]. Hence, increasing the size of abscess definition seems to be a reasonable recommendation.

This study has potential limitations. First, CT image acquisition was not standardized and, depending on additional clinical indications, a wide range of scan protocols was used. However, despite being a frequent examination, until now there is no consensus regarding the scan protocol, and consistent recommendations on how to perform CT examinations in patients with suspected DD are still lacking [9]. Second, the observers had different levels of expertise. However, these different levels reflect daily practice and, most importantly, the agreement scores were comparable. Third, only one reader of each education level was recruited, so we did not perform a subgroup analysis of intra- and interobserver agreement among readers with comparable levels of experience. However, the results of our study demonstrate a high reliability, even in the case of different education levels. Therefore, interobserver agreement among specialists or residents should be less important.

Conclusion

The Classification of Diverticular Disease is a feasible, easy-to-use classification, which can be readily applied in the clinical routine by radiologists with different levels of experience. The CDD has high-grade intra- and interobserver agreement. In particular, it allows the differentiation of acute uncomplicated and complicated diverticulitis, which is crucial in the context of stage-adjusted therapy and prognosis.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Bharucha AE, Parthasarathy G, Ditah I et al. Temporal Trends in the Incidence and Natural History of Diverticulitis: A Population-Based Study. Am J Gastroenterol 2015; 110: 1589–1596

[2] Rustom LBO, Sharara AI. The Natural History of Colonic Diverticulosis: Much Ado about Nothing? Inflamm Intest Dis 2018; 3: 69–74

[3] Strate LL, Modi R, Cohen E et al. Diverticular disease as a chronic illness: evolving epidemiologic and clinical insights. Am J Gastroenterol 2012; 107: 1486–1493

[4] Shahedi K, Fuller G, Bolus R et al. Long-term risk of acute diverticulitis among patients with incidental diverticulosis found during colonoscopy. Clin Gastroenterol Hepatol 2013; 11: 1609–1613

[5] Laurell H, Hansson LE, Gunnarsson U. Acute diverticulitis-clinical presentation and differential diagnostics. Colorectal Dis 2007; 9: 496–501; discussion 501–502

[6] Hupfeld L, Pommergaard HC, Burcharth J et al. Emergency admissions for complicated colonic diverticulitis are increasing: a nationwide register-based cohort study. Int J Colorectal Dis 2018; 33: 879–886

[7] Nagata N, Ishii N, Manabe N et al. Guidelines for Colonic Diverticular Bleeding and Colonic Diverticulitis: Japan Gastroenterological Association. Digestion 2019; 99 (Suppl. 1): 1–26

[8] Leifeld L, Germer CT, Böhm S et al. S3-Leitlinie Divertikelkrankheit/Divertikulitis – Gemeinsame Leitlinie der Deutschen Gesellschaft für Gastroenterologie, Verdauungs- und Stoffwechselkrankheiten (DGVS) und der Deutschen Gesellschaft für Allgemein- und Viszeralchirurgie (DGAV). Z Gastroenterol 2022; 60: 613–688

[9] Meyer S, Schmidbauer M, Wacker FK et al. To Fill or Not to Fill? – Value of the Administration of Positive Rectal Contrast for CT Evaluation of Diverticular Disease of the Colon. Fortschr Röntgenstr 2021; 193: 804–812

[10] Laméris W, van Randen A, Bipat S et al. Graded compression ultrasonography and computed tomography in acute colonic diverticulitis: meta-analysis of test accuracy. Eur Radiol 2008; 18: 2498–2511

[11] Hinchey EJ, Schaal PG, Richards GK. Treatment of perforated diverticular disease of the colon. Adv Surg 1978; 12: 85–109

[12] Ambrosetti P, Grossholz M, Becker C et al. Computed tomography in acute left colonic diverticulitis. Br J Surg 1997; 84: 532–534

[13] Dharmarajan S, Hunt SR, Birnbaum EH et al. The efficacy of nonoperative management of acute complicated diverticulitis. Dis Colon Rectum 2011; 54: 663–671

[14] Sartelli M, Weber DG, Kluger Y et al. 2020 update of the WSES guidelines for the management of acute colonic diverticulitis in the emergency setting. World J Emerg Surg 2020; 15: 32

[15] Klarenbeek BR, de Korte N, van der Peet DL et al. Review of current classifications for diverticular disease and a translation into clinical practice. Int J Colorectal Dis 2012; 27: 207–214

[16] Galetin T, Galetin A, Vestweber KH et al. Systematic review and comparison of national and international guidelines on diverticular disease. Int J Colorectal Dis 2018; 33: 261–272

[17] Schreyer AG, Layer G. S2k Guidelines for Diverticular Disease and Diverticulitis: Diagnosis, Classification, and Therapy for the Radiologist. Fortschr Röntenstr 2015; 187: 676–684

[18] Rausch VH, Weinrich JM, Schön G et al. Accuracy of preoperative CT staging of acute colonic diverticulitis using the classification of diverticular disease (CDD) – Is there a beneficial impact of water enema and visceral obesity? Eur J Radiol 2021; 141: 109813

[19] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159–174

[20] Mayer H et al. Qualitätskriterien von Assessmentinstrumenten – Cohen's Kappa als Maß der Interrater-Reliabilität (Teil 1). Pflege 2004; 17: 36–46

[21] Ünlü Ç, Beenen LFM, Fauquenot JMB et al. Inter-observer reliability of computed tomographic classifications of diverticulitis. Colorectal Dis 2014; 16: O212–O219

[22] Lock JF, Galata C, Reiflfelder C et al. The Indications for and Timing of Surgery for Diverticular Disease. Dtsch Arztebl Int 2020; 117: 591–596

[23] Rao PM, Rhea JT, Novelline RA et al. Helical CT with only colonic contrast material for diagnosing diverticulitis: prospective evaluation of 150 patients. Am J Roentgenol 1998; 170: 1445–1449

[24] Maniatis V, Chryssikopoulos H, Roussakis A et al. Perforation of the alimentary tract: evaluation with computed tomography. Abdom Imaging 2000; 25: 373–379

[25] Lauscher JC, Lock JF, Aschenbrenner K et al. Validation of the German Classification of Diverticular Disease (VADIS) – a prospective bicentric observational study. Int J Colorectal Dis 2021; 36: 103–115